Geophysical Journal International

Geophys. J. Int. (2015) **202,** 347–360
GJI Seismology

doi: 10.1093/gji/ggv138

A comparison of methods to estimate seismic phase delays: numerical examples for coda wave interferometry

T. Dylan Mikesell,[1] Alison E. Malcolm,[1,2] Di Yang[1,3] and Matthew M. Haney[4]

[1] *Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA* 02139, *USA. E-mail: mikesell@mit.edu*
[2] *Memorial University of Newfoundland, St. John's, NL, Canada*
[3] *Exxon-Mobil, Houston, TX, USA*
[4] *Alaska Volcano Observatory, U.S. Geological Survey Volcano Science Center, Anchorage, AK, USA*

Accepted 2015 March 20. Received 2015 March 16; in original form 2014 October 6

SUMMARY

Time-shift estimation between arrivals in two seismic traces before and after a velocity perturbation is a crucial step in many seismic methods. The accuracy of the estimated velocity perturbation location and amplitude depend on this time shift. Windowed cross-correlation and trace stretching are two techniques commonly used to estimate local time shifts in seismic signals. In the work presented here we implement Dynamic Time Warping (DTW) to estimate the warping function – a vector of local time shifts that globally minimizes the misfit between two seismic traces. We compare all three methods using acoustic numerical experiments. We show that DTW is comparable to or better than the other two methods when the velocity perturbation is homogeneous and the signal-to-noise ratio is high. When the signal-to-noise ratio is low, we find that DTW and windowed cross-correlation are more accurate than the stretching method. Finally, we show that the DTW algorithm has good time resolution when identifying small differences in the seismic traces for a model with an isolated velocity perturbation. These results impact current methods that utilize not only time shifts between (multiply) scattered waves, but also amplitude and decoherence measurements.

Key words: Time-series analysis; Interferometry; Seismic monitoring and test-ban treaty verification; Coda waves.

1 INTRODUCTION

Measuring the time shift (or alignment) between two events in a pair of seismic traces is an important step in many geophysical applications. For example, one can use local cross-correlation of neighbouring seismic traces to estimate residual statics (e.g. chapter 3.3 in Yilmaz 2001). In imaging and tomography, one often compares time shifts between observed and synthetic data in order to update the next iteration of a velocity model (e.g. Burdick *et al.* 2013). In reservoir monitoring studies, one looks for time-dependent velocity changes in the subsurface, usually manifested as changes in arrival times (e.g. Khatiwada *et al.* 2012). Therefore, we need robust methods to estimate small arrival time differences between seismic traces, even when the signal-to-noise ratio approaches as low as one. Here we focus on one specific application—coda wave interferometry—but we emphasize that our findings likely have value in other geophysical applications that seek to measure shifts between regularly sampled data either in space or time.

In seismology, the (multiply) scattered wavefield is commonly called the *coda*. Extracting information about the subsurface from this often chaotic looking coda is a difficult task. For decades seismologists have worked to unravel the coda signal, hoping to learn something about either the seismic source or the (an)elastic properties of the Earth's crust. For example, coda amplitudes have been used to determine earthquake magnitude (see chapter 3.2 in Sato *et al.* 2012, and references therein) and forecast volcanic eruptions (Aki & Ferrazzini 2000). Within the last decade, due to the proliferation of the ambient noise correlation method (Shapiro & Campillo 2004; Sabra *et al.* 2005), coda traveltimes have been used as a tool to monitor small velocity changes in the subsurface continuously through time (see Poupinet *et al.* 2008, and references therein). As coda comprises (multiply) scattered waves, these waves often traverse the Earth along a much longer path than the ballistic waves used in tomographic imaging (e.g. Nolet 2008). The fact that the coda have spent more time sampling the subsurface means that they are more sensitive to any velocity perturbations. Therefore, numerous researchers have spent time developing methods to measure velocity changes using coda and then image those changes in space (e.g. Sens-Schönfelder & Wegler 2006; Brenguier *et al.* 2008; Wegler *et al.* 2009; Brenguier *et al.* 2014).

In its most simple form, coda wave interferometry maps coda arrival time perturbation measurements (τ) to velocity perturbations (ΔV) using

$$\frac{\tau}{t} = -\frac{\Delta V}{V}, \tag{1}$$

where t is the lapse time in the coda at which lag τ is measured and V is the background velocity without the perturbation (e.g. Grêt *et al.* 2006, equation 1). We note that this equation is true under the assumption that the velocity perturbation occurs everywhere and with the same amplitude (i.e. ΔV is homogeneous). In this case, we see that there exists a one-to-one mapping of the measured coda arrival time perturbation to the estimated velocity perturbation. Pacheco & Snieder (2005) developed a framework for the spatial and temporal sensitivity of τ when ΔV is not homogeneous. In this case, there is a kernel K that modifies the right hand side of eq. (1). Many authors continue to work on analytical and numerical methods to approximate and compute these sensitivity kernels (Pacheco & Snieder 2005, 2006; Obermann *et al.* 2013; Mayor *et al.* 2014; Planès *et al.* 2014, 2015; Richter *et al.* 2014), but any further treatment of this part of the problem is beyond the scope of this paper.

The remainder of this paper focuses on our ability measure τ. It is obvious that even without the assumption of a homogeneous velocity perturbation, accurate estimation of τ determines the accuracy of the estimated velocity perturbation. In this paper, we consider three time-domain methods to measure τ: (1) windowed cross-correlation, (2) trace stretching and (3) dynamic time warping (DTW). Analogous to windowed cross-correlation is the Fourier domain moving-window cross spectral technique (Poupinet *et al.* 1984), and recent studies have characterized the accuracy of τ estimation from the windowed cross-correlation and stretching methods (e.g. Clarke *et al.* 2011; Zhan *et al.* 2013). In our study we introduce DTW as a new method for coda wave arrival time difference measurements. In doing so, we compare this new approach with existing methods. In the following sections we compare and contrast the three methods using a number of numerical experiments and note that Kanu & Munoz (2014) performed a similar study, comparing these three methods using microseismic event models. We investigate the underlying assumptions in each method and discuss the implications of these assumptions and consequences of each method on the measured τ. In certain cases we find that DTW offers advantages over present methods.

2 THEORY AND COMPARISON OF APPROACHES

In this section we present an introduction to each of the three methods used to measure coda time shifts—(1) windowed cross-correlation, (2) trace stretching, and (3) DTW. We describe the underlying mathematics and assumptions in each of the methods and compare the results of each method using a multiple scattering numerical experiment with a homogeneous velocity perturbation.

2.1 The numerical experiment

We generate synthetic waveforms using the *Seismic Unix* code *sufdmod*. This is a 2-D acoustic finite-difference code. We use a 32×32 km acoustic domain with absorbing boundaries on all four sides. We place a source and a receiver at the centre of the domain and model the acoustic wavefield for 10 s. The source is a Ricker wavelet with a dominant frequency of 12.5 Hz. To speed the time-domain measurement process, waveforms are resampled to a sample interval of 1 ms after the simulations. In the following examples we compare two simulations $u_0(t)$ and $u(t)$.

In the first simulation $u_0(t)$, the background velocity model is $6\,\mathrm{km\,s^{-1}}$ and in the second simulation $u(t)$, the background velocity is $5.94\,\mathrm{km\,s^{-1}}$. The second background velocity model is 1 per cent slower than the first, representing a -1 per cent change (i.e. $\Delta V/V = -1$ per cent). The first background model is shown in Fig. 1(a). The density in each model is the same and we change only the velocity. In order to have phase arrivals to compare arrival time measurements, we need a velocity model with scattering. Therefore, we generate a single zero-mean random Gaussian perturbation model with a perturbation correlation length of 400 m. The random velocity model is shown in Fig. 1(b). Note that the black triangle and star in Figs 1(a) and (b) represent the collocated receiver and source positions, respectively.

We drape this random model over the two different background models and use these models as the input velocity field for the finite-difference simulations. Fig. 1(c) shows the final velocity model used to simulate $u_0(t)$. This type of random velocity model causes scattering of the acoustic waves (Fig. 2), which we use to compare the arrival time difference measurement methods.

2.2 Windowed cross-correlation

The first method we consider is windowed cross-correlation. Arrival time differences have been estimated using windowed cross-correlation (or moving-window cross spectral method in the frequency domain) by numerous authors (e.g. Poupinet *et al.* 1984; Roberts *et al.* 1992; Grêt *et al.* 2006; Snieder 2006; Haney *et al.* 2009). The windowed cross-correlation that we implement is

$$CC(\tau) = \frac{\int_{t_1}^{t_2} u_0(t)u(t+\tau)\,\mathrm{d}t}{\sqrt{\int_{t_1}^{t_2} u_0(t)\,\mathrm{d}t \int_{t_1}^{t_2} u(t)\,\mathrm{d}t}}. \quad (2)$$

This method is referred to as windowed cross-correlation because of the limits t_1 and t_2. In the synthetic example we show in Fig. 3, we arbitrarily set $t_1 = 1$ s and $t_2 = 2$ s. The two input waveforms are shown in the inset of Fig. 2, and the resulting cross-correlation function $CC(\tau)$ is shown in Fig. 3(a).

In this method, the lag $\tau_{\max}$ associated with the cross-correlation maximum $CC_{\max}$ is the value we take to estimate the velocity perturbation. Taking the centre of the analysis window at $t_{\mathrm{mid}} = t_1 + (t_2 - t_1)/2$, we compute the homogeneous velocity perturbation following eq. (1) as

$$\tau_{\max}/t_{\mathrm{mid}} = (0.01/1.5) \times 100 = 0.66\,\text{per cent} = -\Delta V/V. \quad (3)$$

This is close to the actual perturbation of $\Delta V = -1.0$ per cent, and we note that we could improve the accuracy by subsampling the time series or subsampling the correlation function and fitting a polynomial to the peak in the correlation function. However, there is a fundamental reason why the estimate is not exact and will never be exact even with the subsampling trick. The underlying assumption in the windowed cross-correlation measurement is that one trace is a time shifted version of the other trace. If this was the case, then the maximum correlation should equal 1.0 and the lag at the maximum correlation coefficient gives the correct time shift. If we look at the two traces more closely (Fig. 2 inset), we immediately see that the arrival time perturbations on the right side of the window are larger than those on the left side. Therefore, it is obvious that the two traces are not simply time shifted versions of each other, and we should not expect this method to give a correct velocity perturbation. As one may guess, we in fact find that the maximum correlation only reaches 0.98 (Fig. 3a).

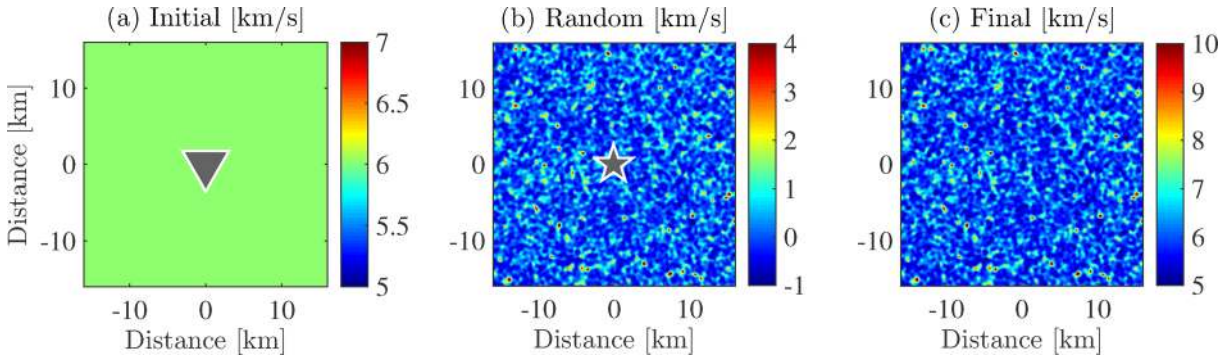

Figure 1. (a) Homogeneous background model; the model domain is 32 × 32 km. (b) Zero-mean random Gaussian model with a correlation length of 400 m. (c) Final velocity model for $u_0(t)$ where (b) is added to (a). The black triangle and star in (a) and (b) are the colocated receiver and source, respectively.

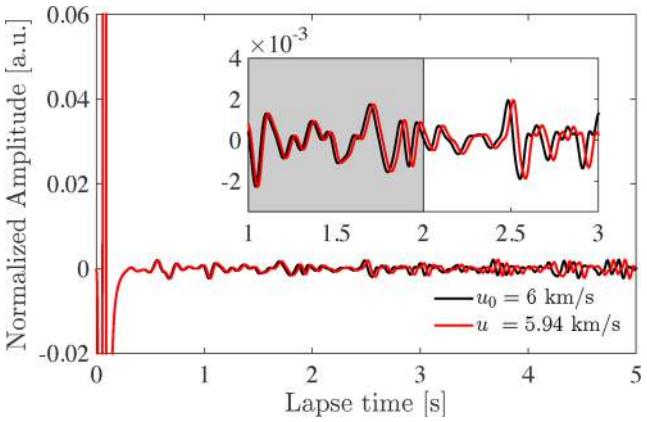

Figure 2. Zero-offset wavefield recorded in the centre of the model domain. The black trace shows the model with background velocity 6 km s^{-1}, which we call u_0. The red trace shows the wavefield for the background velocity model that is 5.94 km s^{-1}, which we call u. Note that the large amplitude direct wave near $t = 0$ is clipped to highlight the scattered waves. Inset: two seconds of the traces; shaded area highlights the time window used to compare methods in Section 2.

2.3 Trace stretching

In the search for improved coda arrival time difference measurements, the trace stretching method was developed (Sens-Schönfelder & Wegler 2006; Hadziioannou *et al.* 2009). The stretching method is based on the assumption that one trace is a stretched version of the other trace, rather than a time shifted version. Furthermore, the stretching is assumed linear with the relationship

$$\tau = \epsilon t, \tag{4}$$

where ϵ is the stretch factor related to the homogeneous velocity perturbation as $\epsilon = -\Delta V/V$. In practice, we stretch the time axis as $t' = t(1 + \epsilon)$. Therefore, a decrease in velocity causes a lengthening of the time axis. After stretching $u(t)$ we then compute the zero-lag cross-correlation over the window $[t_1, t_2]$:

$$CC(\epsilon) = \frac{\int_{t_1}^{t_2} u_0(t)u(t')\,\mathrm{d}t}{\sqrt{\int_{t_1}^{t_2} u_0(t)\,\mathrm{d}t \int_{t_1}^{t_2} u(t')\,\mathrm{d}t}}. \tag{5}$$

To estimate the optimal stretch factor we test many ϵ values and search for the maximum zero-lag cross-correlation coefficient. The stretch function $CC(\epsilon)$ for the trace window from $t_1 = 1$ s to $t_2 = 2$ s is shown in Fig. 3(b).

In this particular case, where we have a homogeneous velocity perturbation, $u(t)$ is actually a stretched version of $u_0(t)$, and we see that we recover a correlation coefficient of 1.0 at exactly $\epsilon = 0.01$. Thus the stretching can be an improvement over the windowed cross-correlation method; however, keep in mind that we have made the assumption that within the analysis window, one trace is a stretched version of the other. Furthermore, as Zhan *et al.* (2013) demonstrated, one must be careful because stretching mixes the phase and amplitude spectra. Therefore, changes in the source amplitude spectra can lead to errors in the stretching coefficient estimation. We will touch on this notion of changing source spectra in Section 6.2.

2.4 Dynamic time warping

The two previously discussed methods both impose assumptions that may not always be valid. In particular, when the velocity perturbation is not homogeneous then only certain parts of the coda are sensitive to the perturbation. In this case, the size and the location of the coda analysis window becomes extremely important as we demonstrate in Section 5. Because of this, we have studied DTW in the seismological context to see how this method works for coda wave arrival-time difference measurement. The name for DTW stems from the field of dynamic programming, which implies breaking one complex problem into many simpler subproblems in order to solve the problem. This idea has been used in seismology before, though maybe not referred to as dynamic programming (e.g. the partitioned waveform inversion of Nolet 1990).

DTW is a non-linear optimization method developed in speech recognition (Sakoe & Chiba 1978; Müller 2007). It has been shown to accurately estimate nonlinear time shifts when strong noise is present (e.g. Hale 2013). Furthermore, DTW has high temporal resolution and suppresses cycle skipping because we fit the entire trace. There is also no averaging over windows. Recent examples of DTW applied to seismology can be found in the seismic exploration literature (e.g. Herrera & van der Baan 2012; Muñoz & Hale 2012; Hale 2013; Yang *et al.* 2014a,b).

If we think of trace stretching as a linear stretch, we can think of DTW as a nonlinear stretch. Consider the seismogram $u_0(t)$, where t is the lapse time. If we apply the time-varying shifts $s(t)$ to $u_0(t)$ such that $t' = t + s(t)$, then we obtain a second seismogram

$$u(t) \approx u_0(t'). \tag{6}$$

In DTW, we want to estimate the shift function $\bar{s}(t)$ that minimizes the misfit between the two traces $u_0(t)$ and $u(t)$. We refer the reader to Hale (2013) for the complete details of this constrained nonlinear

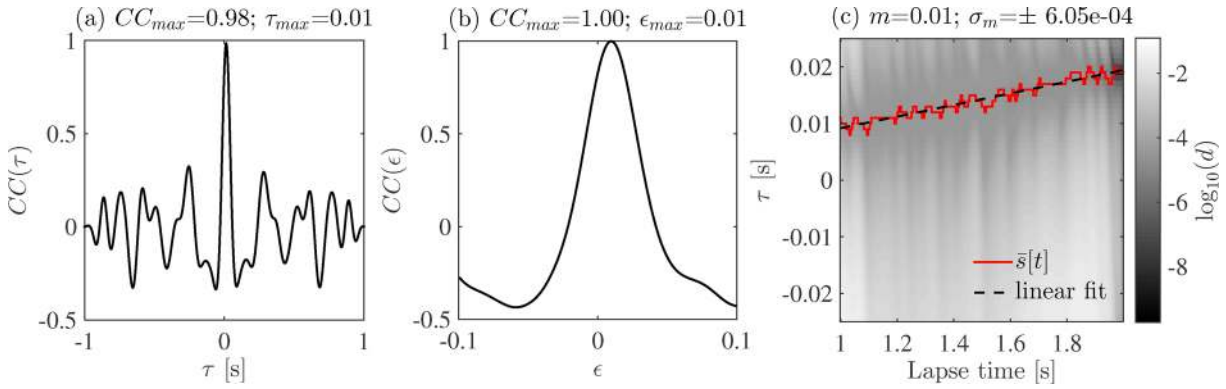

Figure 3. (a) The cross-correlation function (eq. 2) for the two traces shown in Fig. 2 inset. (b) The stretching correlation coefficient (eq. 5). (c) The distance function d (eq. 10), with $b = 5$, the warping function (red), and the linear fit (dashed-black). The linear fit gives a slope of $m = 0.01$.

optimization problem and the boundary conditions, but below we briefly outline the main steps in the algorithm in order to illustrate how the constraints relate to the problem at hand.

The first step in DTW is to compute an error function between the two traces:

$$e(t, \tau) = (u(t) - u_0(t + \tau))^2 , \tag{7}$$

where τ is a vector of lags, defined by the user from $-\tau_{\max}$ to $+\tau_{\max}$ with the same sample interval as the signals ($\mathrm{d}t = d\tau$). Here, we define the error function as the squared difference of amplitudes, but one could also use another norm, such as the absolute value of the difference. The second step is to accumulate these errors in time; thus creating the so-called distance function $d(t, \tau)$:

$$\begin{aligned} d(t_1, \tau) &= e(t_1, \tau), \\ d(t_i, \tau_j) &= e(t_i, \tau_j) + \min \begin{cases} d(t_{i-1}, \tau_{j-1}) \\ d(t_{i-1}, \tau_j) \\ d(t_{i-1}, \tau_{j+1}) \end{cases} , \end{aligned} \tag{8}$$

for $i = 2, 3, \ldots, N$, where N is the number of samples in the two traces and $j = 2, 3, \ldots, M - 1$, where M is the number of elements in the τ vector.

From the distance function we can find the *minimum accumulated distance* D at the last time sample N

$$D = \min_{\tau} d(t_N, \tau). \tag{9}$$

Given that we know D and at which τ D exists, the previous minimum distance must come from a distance on the right hand side of eq. (8). We can then backtrack to recursively find the optimal warping path, from here on called the warping function $\bar{s}(t)$. This is the function that globally minimizes the distance between the two traces (i.e. D), in essence the optimal fit.

Using this approach, $\bar{s}(t)$ is the set of lags that give the global minimum misfit between $u_0(t)$ and $u(t)$ subject to the constraints:

(i) The maximum comparison distance is $\pm\tau_{\max}$

(ii) The maximum step is $\mathrm{d}t$, where $\mathrm{d}t$ is the sample interval (i.e. $|\bar{s}(t_i) - \bar{s}(t_{i-1})| \leq \mathrm{d}t$).

This latter constraint is due to our chosen step pattern defined in eq. (8), and next we discuss modifications to the step pattern that make this optimization better suited to the coda wave interferometry problem.

2.4.1 *Step pattern design*

In the context of velocity perturbations, the step pattern can be thought of as limiting the magnitude of the velocity perturbation. Using eq. (8), the velocity perturbation can be ± 100 per cent, meaning in one *time-step* $\mathrm{d}t$, we can make one *lag-step* $d\tau$. In many geophysical problems this is an unrealistic change, and therefore, we can modify the *time-step* pattern defined in eq. (8) so that maximum perturbation is much less than 100 per cent. Müller (2007) gives example step patterns and the associated boundary conditions in Section 4.2.1. Here we follow Hale (2013) and modify eq. (8) so that

$$\begin{aligned} d(t_1, \tau) &= e(t_1, \tau), \\ d(t_i, \tau_j) &= e(t_i, \tau_j) + \min \begin{cases} d(t_{i-b}, \tau_{j-1}) + \sum_{c=i-1}^{i-b+1} e(t_c, \tau_{j-1}) \\ d(t_{i-1}, \tau_j) \\ d(t_{i-b}, \tau_{j+1}) + \sum_{c=i-1}^{i-b+1} e(t_c, \tau_{j+1}) \end{cases} , \end{aligned} \tag{10}$$

where we have introduced the integer parameter $b \geq 2$ and $i = b + 1$, $b + 2, \ldots, N$ and j is the same as in eq. (8). The DTW optimization now has $d\tau = b * \mathrm{d}t$ and is constrained such that

$$|\bar{s}(t_i) - \bar{s}(t_{i-1})| \leq \frac{d\tau}{b}. \tag{11}$$

This has the effect of limiting the maximum velocity perturbation; for instance if $b = 2$, the maximum velocity perturbation is 50 per cent. This is sometimes referred to in the literature as the slope constraint. For a review of various step patterns we refer the reader to the speech recognition literature (e.g. Rabiner & Juang 1993; Müller 2007), where they also describe boundary conditions for each step pattern.

For the traces in Fig. 2, the $\log_{10}$ of the distance function is shown in Fig. 3(c) with $\tau_{\max}$ set to 0.025 s and $b = 5$ (max $\Delta V/V = 20$ per cent). The red line shows the estimated warping function. Because we are dealing with a homogeneous velocity perturbation $\tau/t = -\Delta V/V$, and we can simply fit a line (dashed-black in Fig. 3 c) to $\bar{s}(t)$ to estimate $\Delta V/V$, which in this case is -1.0 per cent. We note that increasing b will decrease the maximum allowed shift, thus making a smoother estimate of $\bar{s}(t)$. We explore the influence of b in the next section.

To summarize our results of this toy model, in Table 1 we list the consequence imposed on the shape of the warping function $s(t)$ by each of the three methods discussed above. Within a given time window $s(t)$ is assumed (1) constant using the windowed

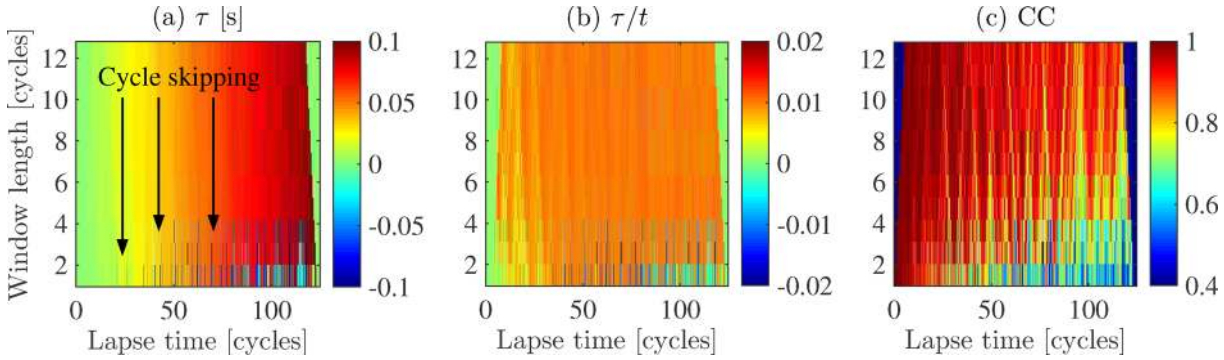

Figure 4. (a) The lag τ at each lapse time for the varying window lengths in the sliding window cross-correlation method. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The maximum correlation coefficient for each window length.

Table 1. Comparison of assumptions imposed on the warping function by each measurement method.

Method	Correlation	Stretching	DTW
$s(t) =$	τ	$t * \epsilon$	$\bar{s}(t)$, subject to $(\Delta V/V)_{\max}$

cross-correlation method, (2) linearly increasing using the stretching method, and (3) any function subject to the constraints already listed using the DTW method.

3 CYCLE SKIPPING

One challenge in coda arrival-time difference measurements, not to mention other methods like waveform inversion, is cycle skipping. The problem of cycle skipping presents itself in many seismic contexts, from waveform inversion to residual statics—really any method that attempts to measure local arrival time differences between waveforms. Cycle skipping in this context is when one wiggle (wave arrival) on the first trace wants to align with the wrong wiggle on the other trace because there are multiple wiggles within the analysis window. This is a common problem if (1) the noise is too large, (2) the analysis window is not long enough, (3) the arrival time difference exceeds the dominant period (e.g. McGuire *et al.* 2012; Richter *et al.* 2014), or (4) the arrivals diverge too much or disappear altogether due to changes in scattering and interference. The cycle skipping problem is further exacerbated when the data are monochromatic or have small bandwidth.

In this section we investigate how well the three measurement methods work when we compare different lapse times in the traces, keeping in mind that as we go to larger lapse times, the time shifts between the wiggles increase. The most common approach to measure local arrival time differences is to slide a window through the two traces, making comparisons in each window. For cross-correlation and stretching, this means that we must choose a window length (e.g. 1 s) and then slide that window (e.g. every 0.5 s) from the beginning to the end of the traces. From each window we collect the maximum correlation value and lag, or ϵ, from windowed cross-correlation or stretching measurements, respectively. Clearly this introduces a parameter (the window length) in the process. In the following subsections we test the windowed cross-correlation and stretching methods using window lengths from 0.1 s to 1.0 s, incrementing at 0.05 s, and we slide these windows through the data one sample at time. When plotting the results we multiply both the window length and lapse time by the dominant source frequency (12.5 Hz) so that the axes are in terms of cycles. This gives a more intuitive feel for how many wiggles must be in a window in order to attain stable measurements.

3.1 Windowed cross-correlation

The lag τ associated with the maximum cross-correlation coefficient for each analysis window and each window length is shown in matrix form in Fig. 4(a). In Fig. 4(b) we plot the lag normalized by lapse time, which gives an estimate of the normalized velocity perturbation $\tau/t = -\Delta V/V$ in each window. We see direct evidence of cycle skipping in Fig. 4(a), indicated by the black arrows. Because we know that all shifts should be positive in this example, when τ goes toward zero or negative time, this indicates cycle skipping, as the wiggles within the analysis window require zero shift or negative shift to align. We note that the occurrence of cycle skipping is directly related to (1) the window length and (2) the lapse time. At short window lengths, the correlation does not find the true shift, and at large lapse times, the shift between the two traces can be larger than the window. Therefore, cycle skipping can continue to exist at large lapse times, even for large window lengths that contain many cycles. This effect can be seen most easily in the correlation coefficient plot in Fig. 4(c), where even at large window lengths the correlation coefficient decreases with lapse time.

Noting the decrease in the maximum cross-correlation coefficient when cycle skipping occurs, in this particular example (i.e., $\Delta V/V = -1$ per cent), we would need a minimum window length of approximately five cycles to prevent cycle skipping at the maximum lapse time. Furthermore, we highlight that along the edges of each panel in Fig. 4 there exists so-called *zero* data that increase in length along the edges from bottom to top. This is because of the t_{mid} variable discussed in Section 2.2. We assign correlation and lag values to the middle of the analysis window; therefore, the edges have no measured data. We could implement a boundary condition that uses a smaller window of data along the edges. However, the measurements would have a different signal-to-noise ratio, and therefore we neglect this in our present analysis.

Before moving on to the stretching method, we modify the windowed cross-correlation method by applying the accumulated lag prior to the cross-correlation. By this, we mean that at window i, we apply τ_{i-1} to $u_0(t_i)$. In this numerical example, this modification reduces the occurrence of cycle skipping (Fig. 5). This is a simple modification that can improve the sliding window cross-correlation results when the signal-to-noise ratio is large. Note that cycle skipping is completely removed and that the colour scale in Fig. 5(c) has a minimum amplitude of 0.95 as compared to 0.4 in Fig. 4(c). There is one caveat to this modification though, and we investigate this in

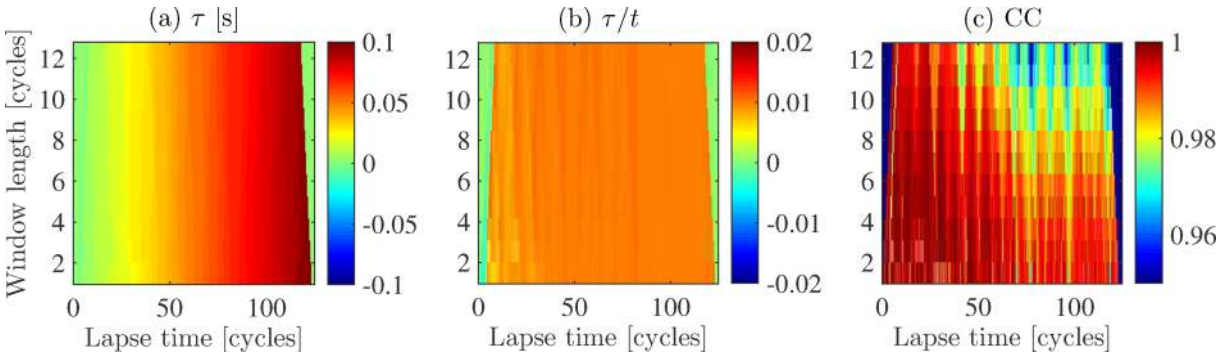

Figure 5. (a) The lag τ at each lapse time for the varying window lengths in the modified sliding window cross-correlation method. Note the absence of cycle skipping when compared to Fig. 4. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The maximum correlation coefficient for each window length.

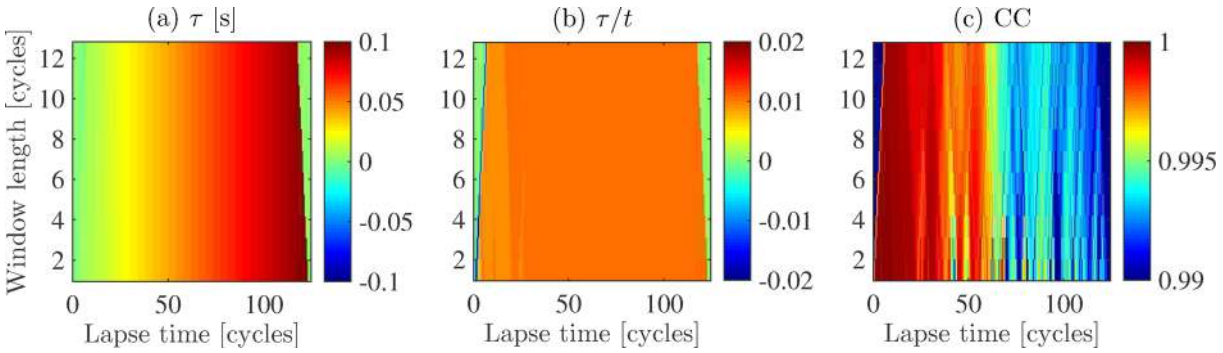

Figure 6. (a) The lag τ at each lapse time for the varying window lengths in the stretching method. Note the absence of cycle skipping when compared to Fig. 4. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The maximum correlation coefficient for each window length.

a later example. The time shift τ_{i-1} must be accurate; otherwise, this modification can degrade τ estimates.

3.2 Stretching

Now we make the same comparison with the stretching method. In Fig. 6 we present the results of applying stretching with the different window lengths. In this case, we see that the τ estimates are even more stable (Fig. 6a) than in the previous correlation examples. We see even less variation in the normalized lag time (Fig. 6b), with the majority of the values estimated very near 1.0 per cent. We also note that there is no cycle skipping present in the stretching method results, and the correlation coefficients [Fig. 6(c)] are even larger than the modified cross-correlation method.

3.3 Dynamic time warping

It is quite obvious that the stretching and the modified cross-correlation methods outperform the traditional windowed cross-correlation method in this particular example. To round out our tests, we show the results for DTW. However, in this case, there is no window size to vary. Therefore, we test various b values and set $\tau_{\max} = 1$ s; we plot in matrix form for comparison with the other methods (Fig. 7). The first thing to note is the consistency between the results for all b values. We see that all b values give extremely similar results, and there is a gradual increase in the residual misfit r (Fig. 7c), which we compute as

$$r = \sum_{i=1}^{N} [u(t_i) - u_0(t_i + \bar{s}(t_i))]^2 . \tag{12}$$

As we restrict the b value we see that the misfit slowly increases, as one might expect when we impose a smooth solution to the optimization problem. The second thing to note is that the edges now have data as DTW analyzes the entire trace, without restrictions from window lengths. We highlight the influence of the direct wave here using Fig. 7(b). DTW does not detect any shift in the direct wave ($\tau/t = 0$), which makes sense as the source and receiver are collocated and the direct wave does not actually propagate through the original or perturbed models.

From this model we can make a few conclusions. We find that the DTW approach gives extremely similar results to the other two methods. There is no cycle skipping present in any of the three methods once we modify the windowed cross-correlation algorithm. Lastly, the edges of the traces are measured using the DTW method. Therefore, DTW appears to be at least as good, if not better, than the two commonly used methods of windowed cross-correlation and stretching for coda wave arrival-time difference estimation in the noise-free case.

What is perhaps more insightful is to look at just how well the different approaches actually estimate the velocity perturbation. To do this we linearly fit the lag as a function of lapse time using weighted ordinary least-squares fitting (Brenguier *et al.* 2008; Haney *et al.* 2015). The weights are the correlation coefficients; therefore, lag windows with low correlation values are down weighted relative to lags with high coefficients in the cross-correlation and stretching

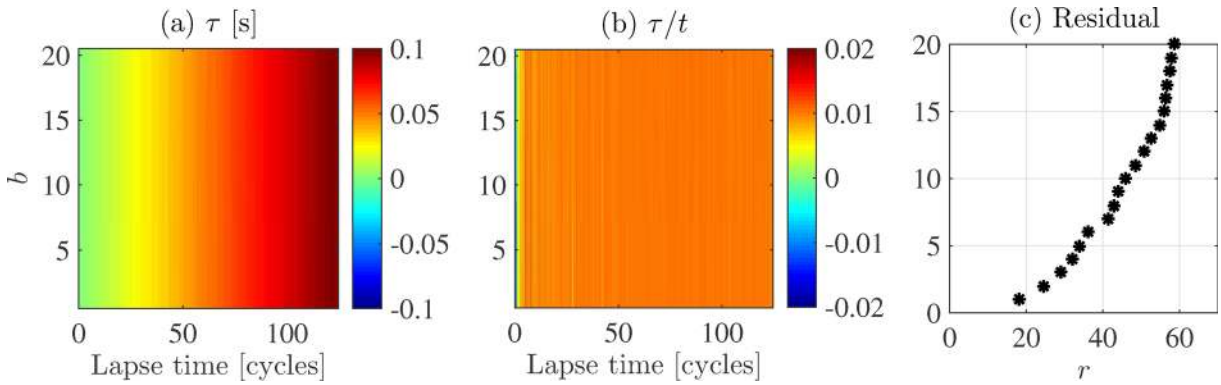

Figure 7. (a) The lag function associated with each b value. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The residual misfit computed along the warping path as defined in eq. (12).

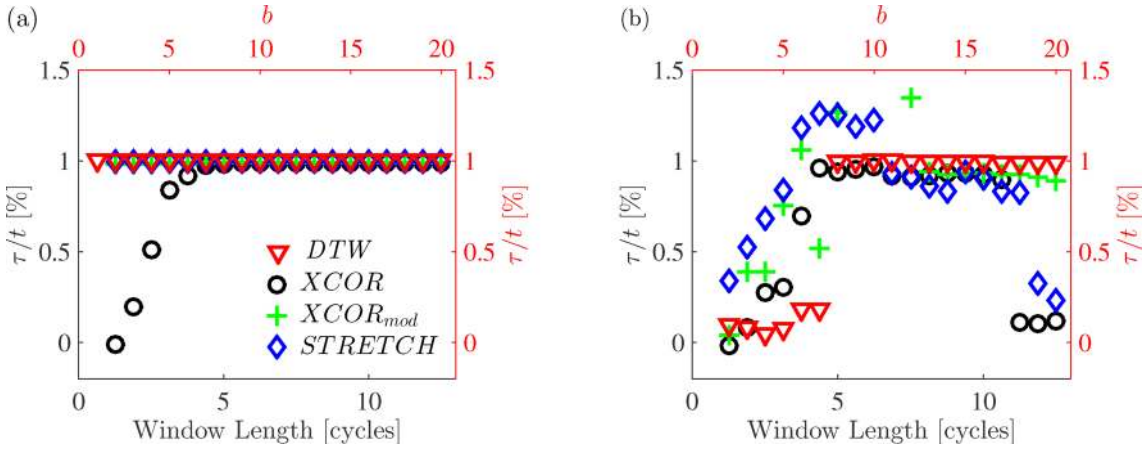

Figure 8. (a) τ/t linear fit for different window lengths (black axes) and b values (red axes) for the noise-free, homogeneous ΔV case. (b) Same as (a) but for the traces with coda SNR = 1. In the windowed cross-correlation and stretching methods we only fit lag times where $CC \geq 0.6$. Legends are the same in both (a) and (b).

methods; all data are weighted equally in the DTW method. We constrain the fitting at $t = 0$ (i.e. $\tau_1 = 0$), and in the cross-correlation and stretching methods, we only use lags with a correlation coefficient of 0.6 or above. Clarke *et al.* (2011) showed that thresholding in this way improves the reliability of velocity perturbation estimates.

We compare the fits for all four methods in Fig. 8(a). For clarity we do not plot the error bars. As we could expect from analyzing the results in Figs 4 to 7, at small window lengths, the accuracy of the fit to the lag function varies, especially for the standard correlation method—the first five points are below 1.0 per cent. As we also would expect, we see a very good agreement between the stretch and DTW methods.

4 LOW SIGNAL-TO-NOISE DATA COMPARISON

To get an idea of how well each method handles low signal-to-noise ratio (SNR) data, we add band-limited (0.5 to 30 Hz) random Gaussian noise to each trace. The noise that we add is zero mean and the amplitude is equal to the maximum amplitude of the coda, where the maximum amplitude of the coda is taken as the maximum absolute value between 2 to 10 s from $u_0(t)$. The SNR of the coda is therefore approximately equal to one. The traces with and without added noise from $t = 1$ s to $t = 3$ s are shown in Fig. 9. Note that the noise is different in each trace.

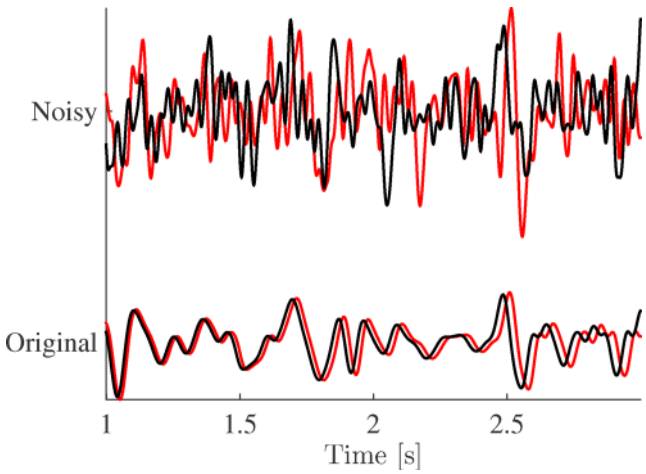

Figure 9. Example of coda from $t = 1$ s to $t = 3$ s with and without band-limited zero mean random Gaussian noise. The SNR of the coda is now approximately equal to one in the Noisy case. Black is u_0 and red is u.

We first compare the two windowed cross-correlation methods; Fig. 10 shows the results from the standard method while Fig. 11 shows the modified method. There is no major difference between the two approaches, and accumulating the lags in the modified method does little to help. If anything, accumulating previous lags may even degrade the results because in some windows we apply

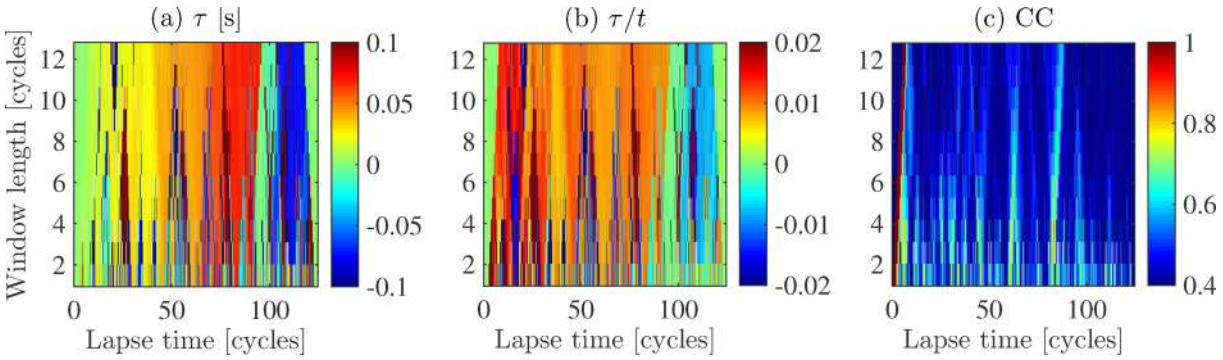

Figure 10. (a) The lag τ at each lapse time for the varying window lengths in the sliding cross-correlation method for noisy data. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The maximum correlation coefficient for each window length.

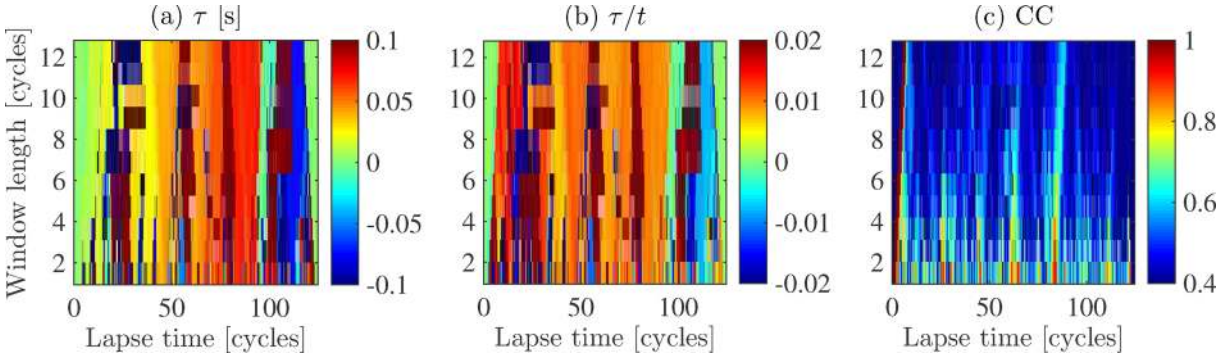

Figure 11. (a) The lag τ at each lapse time for the varying window lengths in the modified sliding cross-correlation method for noisy data. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The maximum correlation coefficient for each window length.

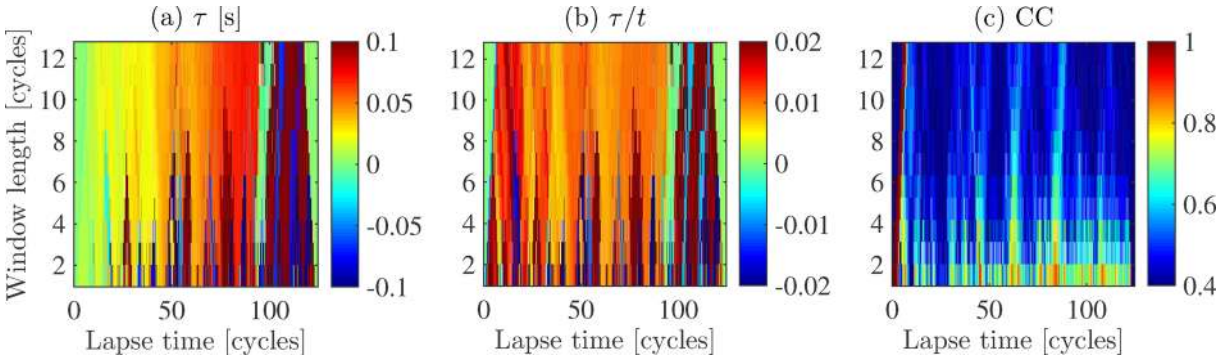

Figure 12. (a) The lag τ at each lapse time for the varying window lengths in the stretching method for noisy data. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The maximum correlation coefficient for each window length.

incorrect lags before correlation. For instance, the modified method overestimates the shift in a number of places around 25 and 50 cycles (indicated by brown patches), regardless of window length. We also find that the cross-correlation coefficients are more similar now between the two methods than in the noise-free case. Lastly, we highlight that both algorithms suffer from cycle skipping and exhibit varying degrees of sensitivity in relation to the window length.

Similar to the windowed cross-correlation, the stretching method does well up to a point (Fig. 12). Around 100 cycles the stretching method also begins to cycle skip—as is evidenced by the blue streaks for all window lengths (Fig. 12a). The noise is too significant at this point in the lapse time. A similar result was found by Hadziioannou *et al.* (2009).

As one might suspect, the DTW results (Fig. 13) are more stable and accurate than the other methods. Regardless of b (except for $1 \leq b \leq 6$), we see that DTW estimates do not suffer from cycle skipping when the SNR is close to one (Fig. 13a); we also see that the estimates of the time shifts are accurate for nearly the entire length of the trace. Finally, in Fig. 13(c) we see a similar variation in the residual misfit given the different b values, and we note that there is little variation in the residual value because the high amplitude direct wave is included in the residual estimate. The residual is larger than for the noise-free case however. Lastly, for now there is no obvious reason as to why $1 \leq b \leq 6$ fails to yield accurate results. We suspect the noise is too large and that we have approached the limits of the DTW method.

To make more quantitative conclusions about the different methods, we again linearly fit the lag functions for each window length and b value with the constraint that $\tau_{i=1} = 0$ (Fig. 8b). For small windows and small b values, all methods give inaccurate estimates of the true perturbations. At larger windows and b values the windowed cross-correlation and DTW methods give estimates near

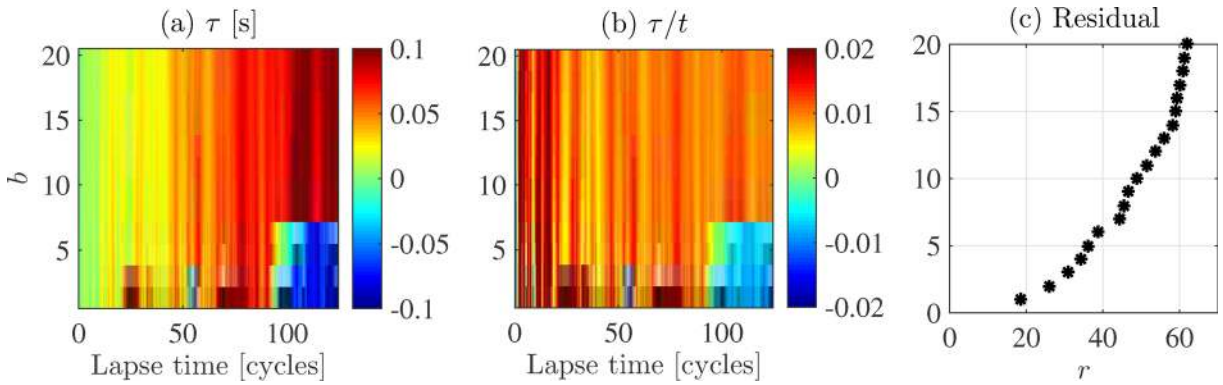

Figure 13. (a) The lag function associated with each b value for noisy data. (b) The time normalized lag giving an estimate of homogeneous negative velocity perturbation. (c) The residual misfit computed along the warping path.

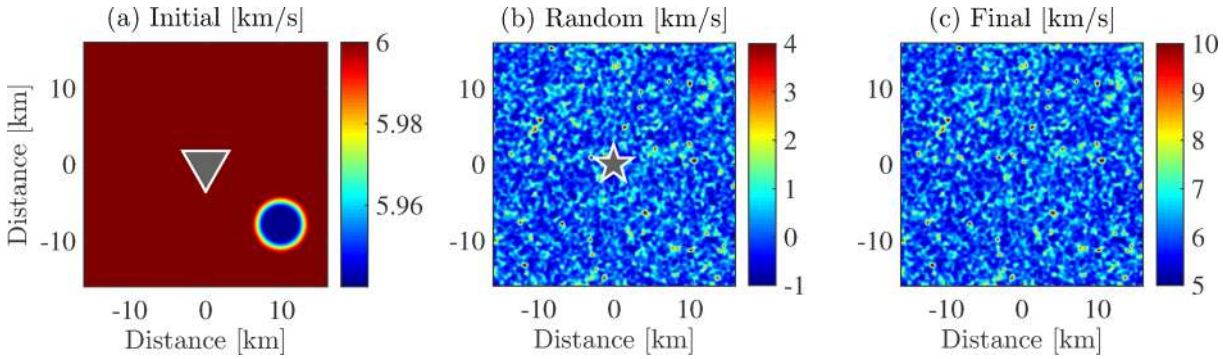

Figure 14. (a) Homogeneous background model with a 1 per cent decrease in the model in the lower right corner; the model domain is 32 × 32 km and the diameter of the perturbation is 4 km. (b) Zero-mean random Gaussian model with a correlation length of 400 m. (c) Final velocity model for $u(t)$ where (b) is added to (a). The black triangle and star in (a) and (b) are the colocated receiver and source, respectively.

1 per cent. The stretching method either over- or underestimates the velocity perturbation. Above $b = 6$, the DTW method estimates the velocity perturbation well; however, all other methods vary as a function of window length. From this we conclude that DTW is a more stable and accurate method to measure homogeneous velocity changes when the data have low SNR.

5 AN ISOLATED VELOCITY PERTURBATION

When the velocity perturbation is heterogeneous, we further violate assumptions in both the windowed cross-correlation and stretching methods. The time shifts are by no means constant, nor are they linearly increasing over a given analysis window. Waves arrive from all parts of the model and likely interfere in time and space with the few waves that have sampled the change. Therefore, one must be clever in choosing the correct (1) window length and (2) lapse time over which to make meaningful measurements. To demonstrate the sensitivity of the methods to this type of isolated velocity perturbation model we again turn to synthetic data.

We make a 1 per cent decrease in velocity in a small circle in the lower right corner of the model domain (Fig. 14a). The transition in the velocity model from the background to the perturbation is a cosine taper function. In numerical testing, a length of 2 km was sufficient to avoid a sharp velocity contrast, which would cause coherent scattering. We want the waves to scatter only due to the random model (Fig. 14b), not the background model. We can see in the complete velocity model (Fig. 14c) that the velocity perturbation is not visually apparent. In this example, $u_0(t)$ is the signal from the model without the perturbation (i.e. $V = 6\,\mathrm{km\,s^{-1}}$ everywhere) and $u(t)$ is the signal with the isolated perturbation in the lower right.

We compare the standard cross-correlation method with the stretching and DTW methods. We plot the lags in Figs 15(a)–(c) for windowed cross-correlation, stretching and DTW, respectively. In Figs 15(d)–(f) we plot the lag-normalized perturbation estimates. It is reasonable to expect localized traveltime changes for an isolated perturbation, and we see that at all window lengths, the cross-correlation and DTW methods are sensitive to the arrival time changes at all times after 50 cycles. However, the stretching method largely detects a single shift around 60 cycles and nothing afterwards. If we compare DTW and windowed cross-correlation to stretching, we realize that the wide orange streak in the latter approach is due to averaging over the analysis window. This leads to the interpretation that all arrival-times in that region have changed, while we can see from DTW and windowed cross-correlation that this is not the case. Interestingly, we also see that the DTW lags show more sensitivity compared to the windowed cross-correlation, where we see reduced amplitudes because of averaging over the correlation window. Investigating DTW to image isolated perturbations will be a future study, but we highlight here that DTW is sensitive to this type of coda wave change.

6 DISCUSSION

We first note that once we have an isolated velocity perturbation, the linear relationship between τ/t and $\Delta V/V$ no longer exists. Instead we need a sensitivity kernel to interpret τ (Pacheco & Snieder 2005, 2006; Planès *et al.* 2014). Therefore, there is little to be gained

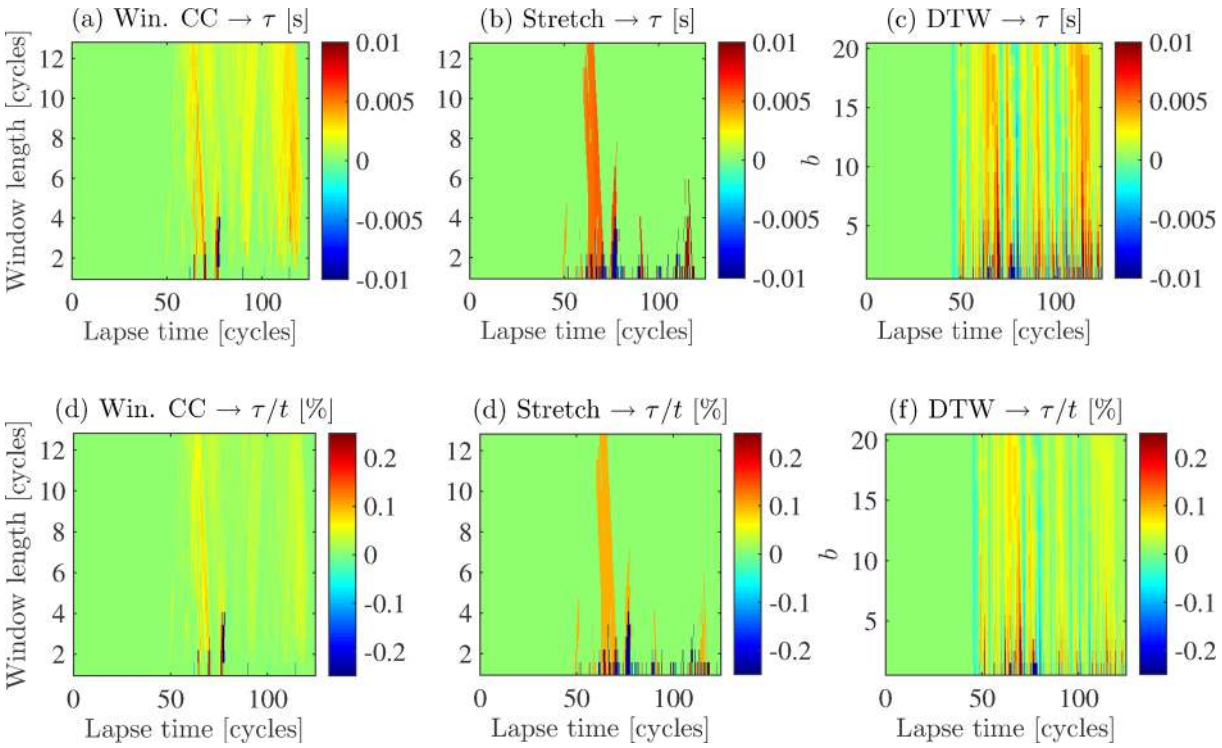

Figure 15. (a–c) Lag τ estimates for isolated velocity perturbations using each method. (d–f) Corresponding normalized time τ/t estimates from each method. The DTW identifies more lapse times with coda changes than do the other two methods.

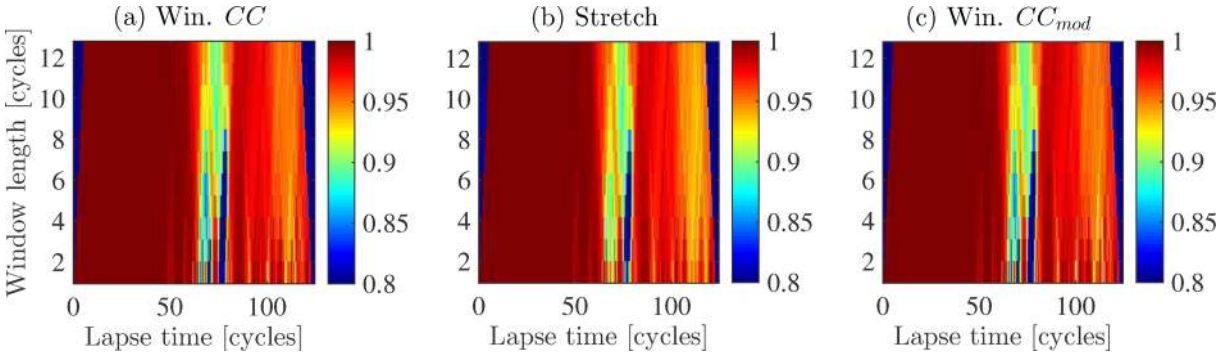

Figure 16. Correlation coefficients for each of the three correlation methods: (a) windowed cross-correlation, (b) stretching and (c) modified windowed cross-correlation. All three identify the changes in the coda beginning around 50 cycles, regardless of the window length.

by comparing lag normalized images in the isolated perturbation example (Figs 15 d–f). To determine which method is more accurate, we would need to actually localize and estimate the amplitude of the velocity perturbation in a tomographic sense (e.g. Obermann *et al.* 2013; Brenguier *et al.* 2014; Planès *et al.* 2014). Instead we consider only τ for the purpose of investigating how well each method does to identify the area of changes. We find that the DTW method improves our capability to locate in time where the changes in coda occur. All of the streaks in Fig. 15(c) match with what the eye visually sees when comparing traces. This is not the case with the stretching method results in Fig. 15(b), which depend heavily on the window size.

In addition to the lag times, one can look at the correlation coefficients. Recent studies have utilized the decorrelation as a means to image localized scatterers (e.g. Larose *et al.* 2010; Rossetto *et al.* 2011; Planès *et al.* 2015). The decorrelation is defined as $1.0 - CC$, where CC is the correlation coefficient. This measurement is related to a combination of the changes in coda scattering and the changes in coda arrival times but is still influenced by window size. The maximum correlation coefficients from the windowed cross-correlation, stretching and modified cross-correlation methods are shown in Fig. 16, respectively. All three correlation methods shown here are sensitive to the (de)correlation. However, we note that we lack these data in the DTW method. Therefore, at the moment it is unclear how DTW might aid this method, but it is worth further investigation as others in seismic exploration have begun studying amplitude warping in combination with time warping (Baek *et al.* 2014).

Finally, there are still aspects to DTW that need to be studied. For one, the decrease in residual as a function of the b value looks like a classic L-curve. It may be of use to study this parameter in the context of Tikhonov regularization (see chapter 5 in Aster *et al.* 2005). Although Hale (2013) notes that b should not be thought of as a regularization parameter, there is an interesting relationship between the shape of the residual curve (e.g. Fig. 13c) and the fluctuations in the τ/t estimates (e.g. Fig. 8b). It would also be interesting

to study DTW in the context of finite bandwidth. Froment *et al.* (2013) demonstrated novel results related to different τ estimates as a function of frequency. Cycle skipping becomes even more severe in narrow band data; DTW may be a useful tool to overcome such problems when analyzing narrow-band data in relation to surface wave depth sensitivity.

6.1 Computation costs

One aspect to consider when using these methods is the computational cost. If we take two traces having N samples and compute the standard time-domain cross-correlation, the cost of this is $O(N^2)$. If we window the two traces such that they have N_1 samples and slide this window one sample at a time through the traces we have a computation cost of $O(N \times N_1^2)$—this is the cost of the windowed cross-correlation method for N windows. In the case of the stretching method the cost can increase significantly. For each ϵ we test we (1) interpolate one trace, (2) window and correlate with the reference trace, and (3) move to the next window and repeat. Fortunately, in this case we are only interested in the zero-lag correlation so the cost of correlation is $O(N_1)$. For M values of epsilon and neglecting the interpolation step, the total cost of stretching is $O(M \times N \times N_1)$. In the DTW algorithm, where we search over K lags, the computation cost is the sum of the three operations: error estimation, accumulation and backtracking. The cost of each is $O(K \times N)$, and the total cost is $O(3 \times K \times N)$.

One can see that depending on M, K, and N_1, the methods have distinct differences. We have neglected the interpolation step, and we note that this can greatly vary the stretching computational costs depending on the choice of interpolation scheme used. We leave it for the user to decide which method is appropriate for their given application, but we note in the study here that DTW and windowed cross-correlation are similar in computation time and that stretching takes much more time. In many cases, sliding the window one sample at a time is overkill and, thus, N may be reduced significantly. Furthermore, depending on the desired ϵ resolution, M may be reduced, and K can be reduced in a similar way.

Lastly, we note two more factors impacting the computation costs. First, the windowed cross-correlation cost can be reduced by implementing the correlation in the frequency domain; the correlation cost is then $O(N_1 \log N_1)$. Secondly, a $b > 1$ value in DTW introduces a multiplication factor $O(b - 1)$ into the accumulation and backtracking steps. For more information on the details of the computation costs with changing step patterns, we refer the reader to chapter 4 in Müller (2007) and the references therein. Specifically, Section 4.3 covering multi-scale DTW may be of use to the readers interested in computationally expensive task such as waveform mining from continuous data.

6.2 Influence of source spectra

It is also worth considering the influence of changing source spectra on the time shift estimates. For a zero-phase wavelet Zhan *et al.* (2013) showed that the stretching method is sensitive to the source spectra of u_0 and u. Changing source spectra between u_0 and u leads to biased velocity perturbation estimates, and they speculate that this is not the case for the windowed cross-correlation method. Therefore, we conduct numerical tests with varying source spectra and compare the three measurement methods. We note a difference between our test and that of Zhan *et al.* (2013). We use a minimum phase wavelet (e.g. chapter 2.2 in Yilmaz 2001), which is by definition casual. Zhan *et al.* (2013) use a zero-phase wavelet and do not make remarks about the causality of their example.

We take the homogeneous background velocity model shown in Fig. 1(c) and compute synthetics for Ricker source wavelets having centre frequencies of 14.0 to 16.0 Hz at a 0.5 Hz interval. The direct wave arrivals for the colocated source and receiver are shown in Fig. 17(a), and the spectra are shown in Fig. 17(b). We analyze the coda wavefield containing many arrivals coming from many different directions and interfering with each other. For comparison with the direct wave, we plot a section of the coda in Fig. 18(a) and the corresponding spectra in Fig. 18(b). The change in the spectrum due to the interfering waves is obvious. This is again different from Zhan *et al.* (2013) who use a single arrival in their analysis. Looking at the time domain wavefields in Figs 17(a) and 18(a), one can see that the delay in the coda arrival follows the same delay as in the direct wave (note the different time scales). Based on these observations of the time-domain signals, we expect that all three methods will estimate spurious velocity changes. Furthermore, if these time delays remain constant over all lapse times, then we will see a decrease in the velocity bias as we go to later lapse times in the coda, as also found by Zhan *et al.* (2013).

In our analysis we take the 14 Hz source wavelet trace as u_0 and compare each of the other traces. The window length in the windowed cross-correlation and stretching techniques is 1.5 s, which is approximately 20 cycles. For the DTW we set $b = 50$, which corresponds to a maximum velocity change of 2 per cent, and we set $\tau_{\max} = 1.0$ s. The lags τ estimated for each of the different source frequencies are shown in Fig. 19. We see that the windowed cross-correlation (a) and DTW (c) give very similar results. They both give more or less a constant lag shift between the different traces. What we see in the stretching method lags is completely different, but actually understood. Looking back at Table 1, we see that the consequence imposed by each method on the warping function $s(t)$ describes what we see in the estimated lags in Fig. 19. The warping function in the windowed cross-correlation method is approximately constant for the different source spectra traces, as is the DTW warping function. The stretch warping function follows a linearly increasing function and then resets giving a sawtooth like function.

If we look at the lag normalized velocity perturbations we can make more interesting comparisons (Fig. 20). For the windowed cross-correlation and DTW methods, we see a pattern in the velocity bias similar to that found by Zhan *et al.* (2013). For the stretching method we see a different pattern that follows the sawtooth pattern in lags. In all three cases, the apparent velocity change decreases with time, with the amplitude of the bias correlated to the difference of the source spectra relative to u_0. We conclude that in the case of a minimum-phase source wavelet, all three methods are susceptible to apparent velocity changes due to differing source spectra between u_0 and u.

7 CONCLUSION

We present a comparison of three methods used to measure small arrival time differences in coda signals caused by weak velocity perturbations. Using an acoustic numerical experiment, we demonstrate the assumptions and consequences inherent in each of the three methods and show that a modification to the windowed cross-correlation algorithm produces results comparable to the stretching algorithm when the SNR is high. Each of the *windowed* methods lacks data at the edges of the signal because each operates over

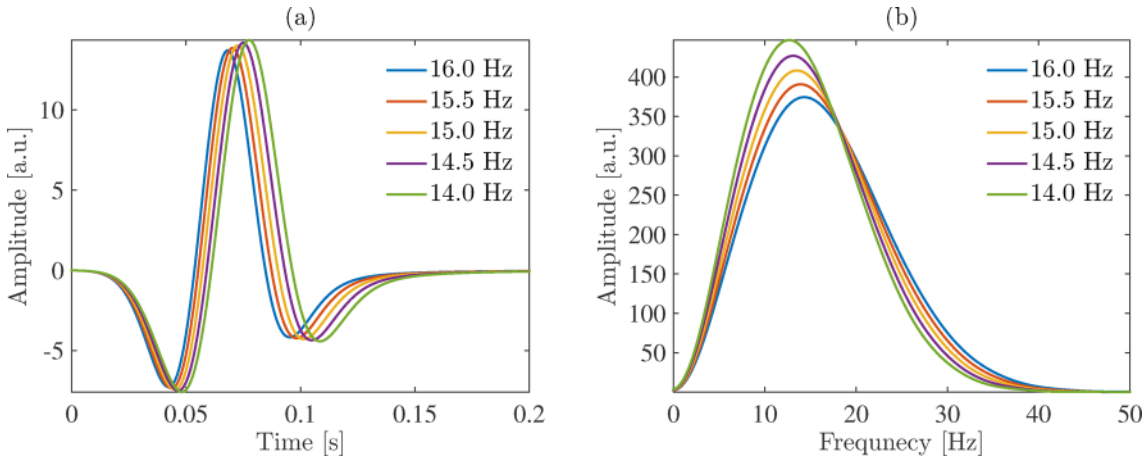

Figure 17. (a) Zero-offset trace displaying the source wavelet. (b) Fourier spectra of traces in (a). The legend denotes the central frequency of the Ricker wavelet, not the dominant frequency.

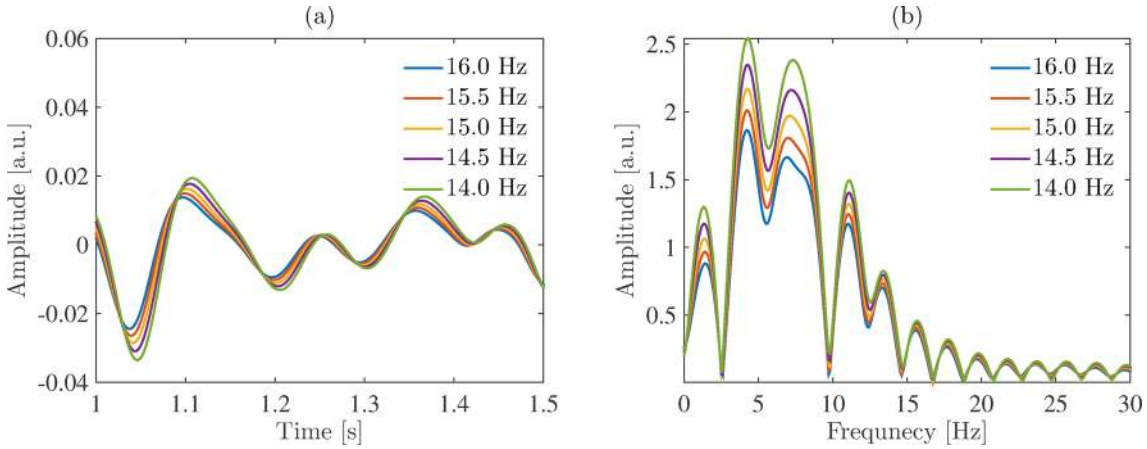

Figure 18. (a) Zero-offset trace displaying the coda from 1 s to 1.5 s. (b) Fourier spectra of traces in (a). The legend denotes the central frequency of the Ricker wavelet, not the dominant frequency.

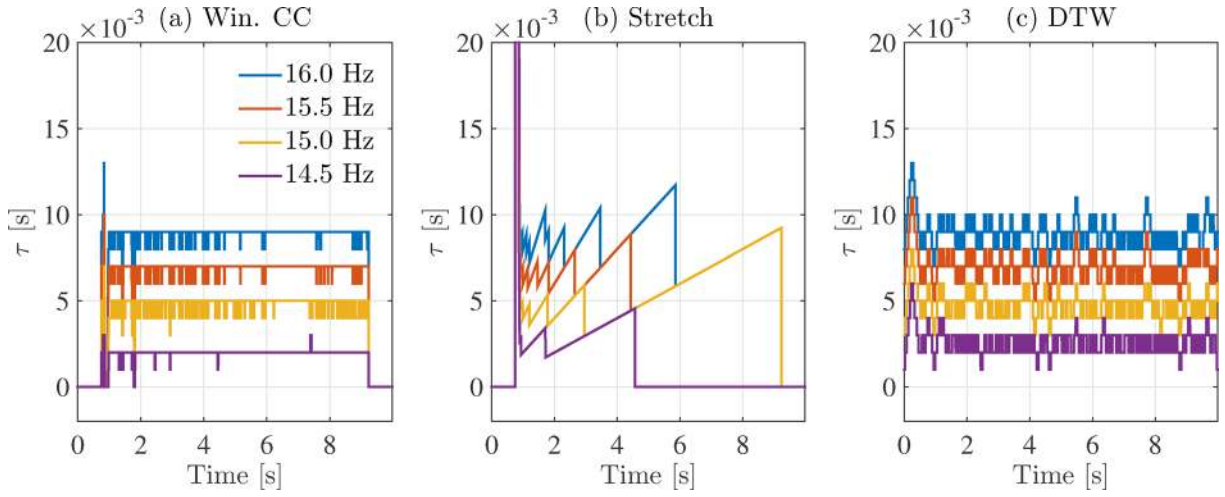

Figure 19. Lags τ measured with each of the three methods: (a) windowed cross-correlation, (b) stretching, and (c) DTW. Legend applies to all plots.

a window of data, and the result is assigned to the centre of that window. We use the DTW method to overcome this lack of data near the edges and suppress cycle skipping. After adding noise to the synthetic data, such that the SNR of the coda was approximately one, we show that DTW is considerably less susceptible to cycle skipping, especially when the arrival time differences become large. Lastly, we demonstrate that when the velocity perturbation is isolated, thus only influencing part of the coda, the DTW method is well suited to determine the location in time of the affected coda. DTW is a new tool that may find new applications in seismology

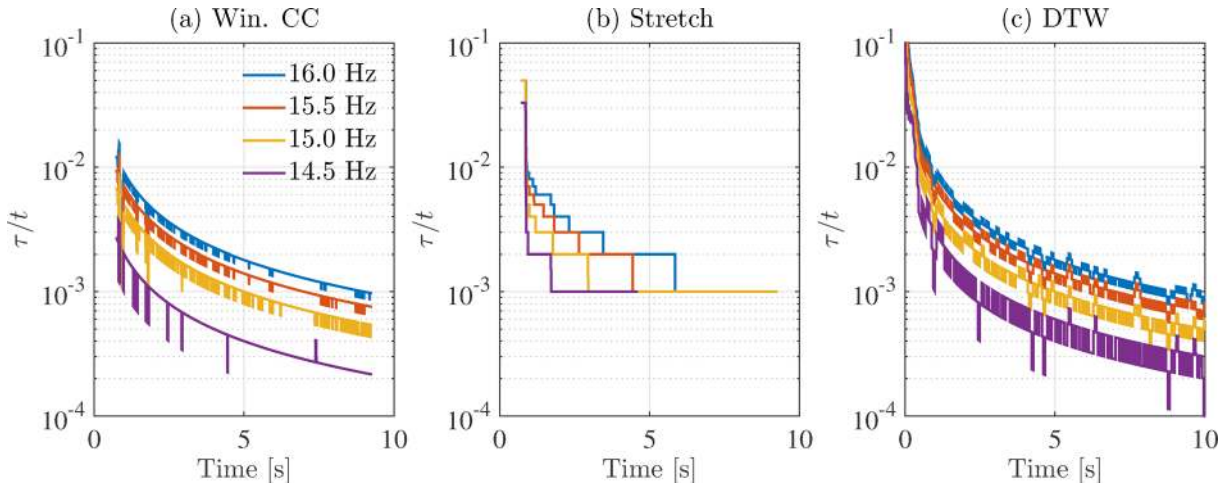

Figure 20. Time normalized lags τ/t measured with each of the three methods: (a) windowed cross-correlation, (b) stretching, and (c) DTW.

and other geophysical methods (e.g. as a waveform inversion misfit function). Future studies will look into choosing the optimal b value and amplitude warping.

ACKNOWLEDGEMENTS

We are grateful to Oleg Poliannikov for providing his code to generate the random velocity model, as well as the developers of Seismic Unix, which we use for the finite difference modelling in this work. We would like to thank Berenice Froment for an example of the stretching method algorithm and Kasper van Wijk and Guust Nolet for fruitful discussions regarding changes in coda waves. TDM acknowledges financial support from the National Science Foundation under Award No. 1144883.

REFERENCES

Aki, K. & Ferrazzini, V., 2000. Seismic monitoring and modeling of an active volcano for prediction, *J. geophys. Res.: Solid Earth,* **105,** doi:10.1029/2000JB900033.

Aster, R., Borchers, B. & Thurber, C.H., 2005. *Parameter Estimation and Inverse Problems,* 1 ed., Elsevier Academic Press.

Baek, H., Burnstad, R. & Keho, T., 2014. Time/Amplitude warping based on multiscale optimization, *2014 SEG Annual Meeting,* pp. 4868–4872.

Brenguier, F., Shapiro, N.M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O. & Nercessian, A., 2008. Towards forecasting volcanic eruptions using seismic noise, *Nature Geoscience,* **1**(2), 126–130.

Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N.M., Briand, X., Emoto, K. & Miyake, H., 2014. Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, *Science,* **345**(6192), 80–82.

Burdick, S., de Hoop, M.V., Wang, S. & van der Hilst, R.D., 2013. Reverse-time migration-based reflection tomography using teleseismic free surface multiples, *Geophys. J. Int.,* **196,** 996–1017.

Clarke, D., Zaccarelli, L., Shapiro, N.M. & Brenguier, F., 2011. Assessment of resolution and accuracy of the Moving Window Cross Spectral technique for monitoring crustal temporal variations using ambient seismic noise, *Geophys. J. Int.,* **186**(1), 867–882.

Froment, B., Campillo, M., Chen, J. & Liu, Q., 2013. Deformation at depth associated with the 12 May 2008 MW 7.9 Wenchuan earthquake from seismic ambient noise monitoring, *Geophys. Res. Lett.,* **40**(1), 78–82.

Grêt, A., Snieder, R. & Scales, J., 2006. Time-lapse monitoring of rock properties with coda wave interferometry, *J. geophys. Res.,* **111,** B03305, doi:10.1029/2004JB003354.

Hadziioannou, C., Larose, E., Coutant, O., Roux, P. & Campillo, M., 2009. Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: laboratory experiments, *J. acoust. Soc. Am.,* **125**(6), 3688–3695.

Hale, D., 2013. Dynamic warping of seismic images, *Geophysics,* **78**(2), S105–S115.

Haney, M.M., van Wijk, K., Preston, L.A. & Aldridge, D.F., 2009. Observation and modeling of source effects in coda wave interferometry at Pavlof volcano, *The Leading Edge,* pp. 554–560.

Haney, M.M., Hotovec-Ellis, A.J., Bennington, N.L., De Angelis, S. & Thurber, C., 2015. Encyclopedia of earthquake engineering, in *Tracking Changes in Volcanic Systems with Seismic Interferometry,* eds Beer, M., Kougioumtzoglou, I.A., Patelli, E. & Au, I.S.-K., Springer, doi:10.1007/978-3-642-36197-5_50-1.

Herrera, R.H. & van der Baan, M., 2012. Automated Seismic-to-well Ties?, *74th EAGE Conference & Exhibition,* p. IO31.

Kanu, C.O. & Munoz, A., 2014. A comparison of three methods for estimating velocity changes between time-lapse microseismic signals, *Tech. rep.,* Colorado School of Mines, CWP, Golden, CO.

Khatiwada, M., Adam, L., Morrison, M. & van Wijk, K., 2012. A feasibility study of time-lapse seismic monitoring of CO_2 sequestration in a layered basalt reservoir, *J. appl. Geophys.,* **82,** 145–152.

Larose, E., Planès, T., Rossetto, V. & Margerin, L., 2010. Locating a small change in a multiple scattering environment, *Appl. Phys. Lett.,* **96,** 204101, http://dx.doi.org/10.1063/1.3431269.

Mayor, J., Margerin, L. & Calvet, M., 2014. Sensitivity of coda waves to spatial variations of absorption and scattering: radiative transfer theory and 2-D examples, *Geophys. J. Int.,* doi:10.1093/gji/ggu046.

McGuire, J.J., Collins, J.A., Gouédard, P., Roland, E., Lizarralde, D., Boettcher, M.S., Behn, M.D. & van der Hilst, R.D., 2012. Variations in earthquake rupture properties along the Gofar transform fault, East Pacific Rise, *Nature Geoscience,* **5**(5), 336–341.

Muñoz, A. & Hale, D., 2012. Automatically tying well logs to seismic data, *Tech. rep.,* CWP-725, Colorado School of Mines, Golden, CO.

Müller, M., 2007. *Information Retrieval for Music and Motion,* Springer, doi:10.1007/978-3-540-74048-3.

Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.: Solid Earth,* **95**(B6), 8499–8512.

Nolet, G., 2008. *A Breviary of Seismic Tomography: Imaging the Interior of the Earth and Sun,* Cambridge University Press.

Obermann, A., Planès, T., Larose, E. & Campillo, M., 2013. Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise, *J. geophys. Res.: Solid Earth,* **118**(12), 1–10.

Pacheco, C. & Snieder, R., 2005. Time-lapse travel time change of multiply scattered acoustic waves, *J. acoust. Soc. Am.,* **118**(3), 1300–1310.

Pacheco, C. & Snieder, R., 2006. Time-lapse traveltime change of singly scattered acoustic waves, *Geophys. J. Int.,* **165**(2), 485–500.

Planès, T., Larose, E., Margerin, L., Rossetto, V. & Sens-Schönfelder, C., 2014. Decorrelation and phase-shift of coda waves induced by local changes: multiple scattering approach and numerical validation, *Waves in Random and Complex Media,* pp. 1–27.

Planès, T., Larose, E., Rossetto, V. & Margerin, L., 2015. Imaging multiple local changes in heterogeneous media with diffuse waves, *J. acoust. Soc. Am.,* **137**(2), 660–667.

Poupinet, G., Ellsworth, W.L. & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. geophys. Res.,* **89**(B7), 5719–5731.

Poupinet, G., Got, J.L. & Brenguier, F., 2008. Chapter 14: monitoring temporal variations of physical properties in the crust by cross-correlating the waveforms of seismic doublets, *Adv. Geophys.,* **50,** 373–399.

Rabiner, L. & Juang, B.-H., 1993. *Fundamentals of Speech Recognition,* Prentice-Hall, Inc.

Richter, T., Sens-Schönfelder, C., Kind, R. & Asch, G., 2014. Comprehensive observation and modeling of earthquake and temperature-related seismic velocity changes in northern Chile with passive image interferometry, *J. geophys. Res.: Solid Earth,* **119,** 4747–4765.

Roberts, P.M., Phillips, W.S. & Fehler, M.C., 1992. Development of the active doublet method for measuring small velocity and attenuation changes in solids, *J. acoust. Soc. Am.,* **91**(6), 3291–3302.

Rossetto, V., Margerin, L., Planès, T. & Larose, E., 2011. Locating a weak change using diffuse waves: theoretical approach and inversion procedure, *J. appl. Phys.,* **109,** 034903, http://dx.doi.org/10.1063/1.3544503.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W.A. & Fehler, M.C., 2005. Extracting time-domain Green's function estimates from ambient seismic noise, *Geophys. Res. Lett.,* **32,** L03310, doi:10.1029/2004GL021862.

Sakoe, H. & Chiba, S., 1978. Dynamic programming algorithm optimization for spoken word recognition, *IEEE Trans. Acoustics, Speech, Signal Processing,* ASSP-26(1), 43–49.

Sato, H., Fehler, M.C. & Maeda, T., 2012. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth,* 2nd ed., Springer.

Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.,* **33**(L21302), doi:10.1029/2006GL027797.

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31**(7), n/a–n/a, doi:10.1029/2004GL019491.

Snieder, R., 2006. The Theory of Coda Wave Interferometry, *Pure appl. Geophys.,* **163,** 455–473.

Wegler, U., Nakahara, H., Sens-Schönfelder, C., Korn, M. & Shiomi, K., 2009. Sudden drop of seismic velocity after the 2004 Mw 6.6 mid-Niigata earthquake, Japan, observed with Passive Image Interferometry, *J. geophys. Res.,* **114,** B06305, doi:10.1029/2008JB005869.

Yang, D., Malcolm, A. & Fehler, K., 2014a. Using image warping for time-lapse image domain wavefield tomography, *Geophysics,* **79**(3), WA141–WA151.

Yang, D., Shang, X., Malcolm, A. & Fehler, M., 2014b. Image Registration Guided Wavefield Tomography For Shear Wave Velocity Model Building, in *SEG Technical Program Expanded Abstracts,* **1,** pp. 1200–1205, Society of Exploration Geophysicists, doi:10.1190/segam2014-0847.1.

Yilmaz, O., 2001. *Seismic Data Analysis,* Society of Exploration Geophysicists.

Zhan, Z., Tsai, V.C. & Clayton, R.W., 2013. Spurious velocity changes caused by temporal variations in ambient noise frequency content, *Geophys. J. Int.,* **194**(3), 1574–1581.